

A LOWER BOUND ON DIMENSION REDUCTION FOR TREES IN ℓ_1

JAMES R. LEE AND MOHAMMAD MOHARRAMI

ABSTRACT. There is a constant $c > 0$ such that for every $\varepsilon \in (0, 1)$ and $n \geq 1/\varepsilon^2$, the following holds. Any mapping from the n -point star metric into ℓ_1^d with bi-Lipschitz distortion $1 + \varepsilon$ requires dimension

$$d \geq \frac{c \log n}{\varepsilon^2 \log(1/\varepsilon)}.$$

1. INTRODUCTION

Consider an integer $n \geq 1$. The n -node star is the simple, undirected graph $G_n = (V_n, E_n)$ with $|V_n| = n$, where one node has degree $n - 1$ and all other nodes have degree one. We write ρ_n for the shortest-path metric on G_n where each edge is equipped with a unit weight. We use ℓ_1^d to denote the space \mathbb{R}^d equipped with the ℓ_1 norm. Our main theorem follows.

Theorem 1. *There is a constant $c > 0$ such that the following holds. Consider any $\varepsilon \in (0, \frac{1}{16})$ and $n \geq 1/\varepsilon^2$. Suppose there exists a 1-Lipschitz mapping $f : V_n \rightarrow \ell_1^d$ such that $\|f(x) - f(y)\|_1 \geq (1 - \varepsilon)\rho_n(x, y)$ for all $x, y \in V_n$. Then,*

$$d \geq \frac{c \log n}{\varepsilon^2 \log(1/\varepsilon)}.$$

One can achieve such a mapping with $d \leq O\left(\frac{\log n}{\varepsilon^2}\right)$, thus the theorem is tight up to the factor of $c/\log(1/\varepsilon)$. In general, de Mesmay and the authors [11] proved that every n -point tree metric admits a distortion $1 + \varepsilon$ embedding into $\ell_1^{C(\varepsilon) \log n}$ where $C(\varepsilon) \leq O((\frac{1}{\varepsilon})^4 \log \frac{1}{\varepsilon})$. For the special case of complete trees where all internal nodes have the same degree (such as the n -star), they achieve $C(\varepsilon) \leq O(\frac{1}{\varepsilon^2})$.

We recall that given two metric spaces (X, d_X) and (Y, d_Y) and a map $f : X \rightarrow Y$, one defines the *Lipschitz constant of f* by

$$\|f\|_{\text{Lip}} = \sup_{x \neq y \in X} \frac{d_Y(f(x), f(y))}{d_X(x, y)}.$$

The *bi-Lipschitz distortion of f* is the quantity $\text{dist}(f) = \|f\|_{\text{Lip}} \cdot \|f^{-1}\|_{\text{Lip}}$, which is taken as infinite when f is not one-to-one. If there exists such a map f with distortion D , we say that X *D -embeds into Y* .

A *finite tree metric* is a finite, graph-theoretic tree $T = (V, E)$, where every edge is equipped with a positive length. The metric d_T on V is given by taking shortest paths. Since every finite tree metric embeds isometrically into ℓ_1 , one can view the preceding statements as quantitative bounds on the dimension required to achieve such an embedding with small distortion (instead of isometrically).

Such questions have a rich history. Perhaps most famously, if X is an n -point subset of ℓ_2 , then a result of Johnson and Lindenstrauss [9] states that X admits a $(1 + \varepsilon)$ -embedding into ℓ_2^d where $d = O\left(\frac{\log n}{\varepsilon^2}\right)$. Alon [2] proved that this is tight up to a $\log(1/\varepsilon)$ factor: If $X \subseteq \ell_2^n$ is an orthonormal basis, then any D -embedding of X into ℓ_1^d requires $d \geq \frac{\Omega(\log n)}{\varepsilon^2 \log(1/\varepsilon)}$.

The situation for finite subsets of ℓ_1 is quite a bit more delicate. Talagrand [18], following earlier results of Bourgain-Lindenstrauss-Milman [5] and Schechtman [17], showed that every n -dimensional subspace $X \subseteq \ell_1$ (and, in particular, every n -point subset) admits a $(1 + \varepsilon)$ -embedding into ℓ_1^d , with $d \leq O\left(\frac{n \log n}{\varepsilon^2}\right)$. For n -point subsets, this was improved to $d \leq O(n/\varepsilon^2)$ by Newman and Rabinovich [15], using the spectral sparsification techniques of Batson, Spielman, and Srivastava [4].

In contrast, Brinkman and Charikar [6] proved that there exist n -point subsets $X \subseteq \ell_1$ such that any D -embedding of X into ℓ_1^d requires $d \geq n^{\Omega(1/D^2)}$ (see also [12] for a simpler argument). Thus the exponential dimension reduction achievable in the ℓ_2 case cannot be matched for the ℓ_1 norm. More recently, it has been shown by Andoni, Charikar, Neiman, and Nguyen [3] that there exist n -point subsets such that any $(1 + \varepsilon)$ -embedding requires dimension at least $n^{1-O(1/\log(\varepsilon^{-1}))}$. Regev [16] has given an elegant proof of both these lower bounds based on information theoretic arguments. Our proof takes some inspiration from Regev's approach.

We note that Theorem 1 has an analog in coding theory. Let $U_n = \{e_1, e_2, \dots, e_n\} \subseteq \ell_1$. Then any $(1 + \varepsilon)$ -embedding of U_n into the Hamming cube $\{0, 1\}^d$ requires $d \geq \frac{\Omega(\log n)}{\varepsilon^2 \log(1/\varepsilon)}$. This was proved in 1977 by McEliece, Rodemich, Rumsey, and Welch [14] using the Delsarte linear programming bound [8]. The corresponding coding question concerns the maximum number of points $x_1, x_2, \dots \in \{0, 1\}^d$ which satisfy $(1 - \varepsilon)d/2 \leq \|x_i - x_j\|_1 \leq (1 + \varepsilon)d/2$ for $i \neq j$. Alon's result for ℓ_2 [2] yields this bound as a special case since $\|x - y\|_2 = \sqrt{\|x - y\|_1}$ when $x, y \in \{0, 1\}^d$.

On the one hand, the lower bound of Theorem 1 is stronger since it applies to the target space ℓ_1^d and not simply $\{0, 1\}^d$. On the other hand, it is somewhat weaker since embedding U_n corresponds to embedding only the leaves of the star graph G_n , while our lower bound requires an embedding of the internal vertex as well. In fact, this is used in a fundamental and crucial way in our proof. Still, in Section 3, we prove the following somewhat weaker lower bound using simply the set U_n .

Theorem 2. *There is a constant $c > 0$ such that for every $\varepsilon \in (0, 1)$, for all n sufficiently large, any $(1 + \varepsilon)$ -embedding of $U_n \subseteq \ell_1^n$ into ℓ_1^d requires*

$$d \geq \frac{c \log n}{\varepsilon \log \frac{1}{\varepsilon}}.$$

For the case of isometric embeddings (i.e., $\varepsilon = 0$), Alon and Pudlák [1] showed that if U_n embeds isometrically in ℓ_1^d , then $d \geq \Omega(n/(\log n))$. Our proof of Theorem 2 bears some similarity to their approach.

Finally, we mention that if B_h denotes the height- h complete binary tree (which has $2^{h+1} - 1$ nodes), then it was proved by Charikar and Sahai [7] that for every $h \geq 1$ and $\varepsilon > 0$, B_h admits a $(1 + \varepsilon)$ -embedding into ℓ_1^d with $d \leq O(h^2/\varepsilon^2)$. It was asked in [13] whether one could achieve $d \leq O(h/\varepsilon^2)$ and this was resolved positively in [11]. From Theorem 1, one can deduce that this upper bound is asymptotically tight up to the familiar factor of $\log(1/\varepsilon)$. This corollary is proved in Section 4.

Corollary 3. *For any $\varepsilon > 0$ and $k \geq 2$, the following holds. For h sufficiently large, any $(1 + \varepsilon)$ -embedding of the complete k -ary, height- h tree into ℓ_1^d requires*

$$d \geq \frac{\Omega(h \log k)}{\varepsilon^2 \log(1/\varepsilon)}.$$

2. PROOF OF THEOREM 1

We will first bound the number of “almost disjoint” probability measures that can be put on a finite set. Then we will translate this to a lower bound for the dimension required for embedding the n -star into ℓ_1^d with distortion $1 + \varepsilon$.

Let X be a finite ground set, and let \mathcal{S} be a set of measures X . We say that \mathcal{S} is ε -unrelated if, for all distinct elements $\mu, \nu \in \mathcal{S}$,

$$\|\mu - \nu\|_{TV} \geq \frac{1}{2}(\mu(X) + \nu(X)) - \varepsilon,$$

where $\|\cdot\|_{TV}$ denotes the total variation distance. The following lemma is an easy corollary of a fact from [16]. We include the proof here for completeness.

Lemma 4. *For every $\varepsilon \in (0, 1)$ and $k \in \mathbb{N}$, if there exists a map $f : (V_n, \rho_n) \rightarrow \ell_1^k$ with distortion $1 + \varepsilon$, then there exists an ε -unrelated set of probability measures on $\{1, \dots, 2k + 1\}$ of size $n - 1$.*

Proof. Let $r \in V_n$ denote the the vertex of degree $n - 1$. By translation and scaling, we may assume that $f(r) = 0$ and f is 1-Lipschiz. Thus for all vertices $v \in V_n$, we have $\|f(v)\|_1 \leq 1$. For each vertex $v \in V_n \setminus \{r\}$ define the measure μ_v as follows

$$\mu_v(\{i\}) = \begin{cases} \max(0, f(v)_i) & 1 \leq i \leq k \\ \max(0, -f(v)_i) & k + 1 \leq i \leq 2k \\ 1 - \|f(v)\|_1 & i = 2k + 1, \end{cases}$$

where we use $f(v)_i$ to denote the i th coordinate of $f(v)$.

Note that for all $u, v \in V_n \setminus \{r\}$ we have

$$\begin{aligned} \|\mu_u - \mu_v\|_{TV} &= \frac{1}{2} \left(\|f(u) - f(v)\|_1 + \left| (1 - \|f(u)\|_1) - (1 - \|f(v)\|_1) \right| \right) \\ &\geq \|f(u) - f(v)\|_1. \end{aligned}$$

Since f has distortion $1 + \varepsilon$, for any two distinct vertices $u, v \in V_n$, we have

$$\|f(u) - f(v)\|_1 \geq \left(\frac{2}{1 + \varepsilon} \right) \geq 2(1 - \varepsilon).$$

Therefore the collection $\{\mu_v : v \in V_n \setminus \{r\}\}$ satisfies the conditions of the lemma. \square

The next lemma is the final ingredient that we need to prove Theorem 1. Let \mathcal{M}_k be the set of all measures $\{1, 2, \dots, k\}$, and let \mathcal{P}_k be the set of all probability measures on $\{1, 2, \dots, k\}$.

Lemma 5. *There exists a universal constant $C \geq 1$ such that for $\varepsilon \leq 1/16$, the following holds. If there is an ε -unrelated set $\mathcal{S} \subseteq \mathcal{P}_k$, then there exists a $\frac{1}{2}$ -unrelated set $\mathcal{T} \subseteq \mathcal{P}_k$ of size at least $\frac{|\mathcal{S}|}{14}$ such that for all $\mu \in \mathcal{T}$, we have $|\text{supp}(\mu)| \leq \lceil C\varepsilon(\varepsilon + \frac{1}{n})d \rceil$.*

Before proving the lemma, we use it to finish the proof of Theorem 1.

Proof of Theorem 1. Suppose that there is a map from the n -star to ℓ_1^d with distortion $1 + \varepsilon$. Then by Lemma 4, there exists an ε -unrelated set of probability measures on $\{2d + 1\}$ of size $n - 1$. Thus by Lemma 5, there must exist a $\frac{1}{2}$ -unrelated set \mathcal{S} of probability measures on $\{1, \dots, 2d + 1\}$ of size $\Omega(n)$ such that every measure in \mathcal{S} has support size at most

$$\left\lceil C \cdot \varepsilon \cdot \left(\varepsilon + \frac{1}{n-1} \right) \cdot (2d + 1) \right\rceil,$$

for some universal constant $C \geq 1$.

We now divide the problem into two cases. In the case that $C\varepsilon(\varepsilon + \frac{1}{|\mathcal{S}|})(2d + 1) < 1$, every measure in \mathcal{S} is supported on exactly one element, therefore $|\mathcal{S}| \leq 2d + 1$. Hence,

$$d \geq \Omega(|\mathcal{S}|) \geq \Omega(n) \geq \frac{\Omega(\log n)}{\varepsilon^2 \log(1/\varepsilon)},$$

where we have used the assumption that $n \geq 1/\varepsilon^2$.

In the second case, we have $C\varepsilon(\varepsilon + \frac{1}{|\mathcal{S}|})(2d + 1) \geq 1$. Since $\frac{1}{|\mathcal{S}|} = O(\varepsilon)$, each element $\mu \in \mathcal{S}$ has $|\text{supp}(\mu)| \leq O(\varepsilon^2 d)$. Thus for some constant $c > 0$, there are at most $\binom{2d+1}{c\varepsilon^2 d} \leq \exp(O(\varepsilon^2 d \log(1/\varepsilon)d))$ different supports of size $O(\varepsilon^2 d)$ for the measures in \mathcal{S} .

Since \mathcal{S} is a $\frac{1}{2}$ -unrelated set of probability measures, for any $\mu, \nu \in \mathcal{S}$, we have

$$\|\mu - \nu\|_{TV} \geq \frac{1}{2}.$$

In particular, if we fix a set $Q \subseteq X$, then by a simple $|Q|$ -dimensional volume argument,

$$|\mu \in \mathcal{S} : \text{supp}(\mu) \subseteq Q| \leq 3^{|Q|}.$$

All together, we have

$$|\mathcal{S}| \leq 3^{O(\varepsilon^2 d)} \cdot e^{O(\varepsilon^2 d \log(1/\varepsilon))} \leq e^{O(\varepsilon^2 d \log(1/\varepsilon))}.$$

Hence, $d \geq \Omega\left(\frac{\log |\mathcal{S}|}{\varepsilon^2 \log(1/\varepsilon)}\right)$, completing the proof. \square

Remark 6. We note that there is a straightforward volume lower bound for large distortions $D \geq 1$: Any D -embedding of the n -star into ℓ_1^d requires $d \geq \Omega\left(\frac{\log n}{\log D}\right)$. This is simply because the maximal number of disjoint ℓ_1 balls of radius $1/D$ that can be packed in an ℓ_1 ball of radius 2 is $(2D)^d$ in d dimensions.

We are left to prove Lemma 5. We start by recalling some simple properties of the total variation distance. For a finite set S and measures $\mu, \nu : 2^S \rightarrow [0, \infty)$, we define

$$\min(\mu, \nu)(T) = \sum_{x \in T} \min\{\mu(\{x\}), \nu(\{x\})\}.$$

For $k \in \mathbb{N}$, and measures $\mu, \nu \in \mathcal{M}_k$, we have

$$\|\mu - \nu\|_{TV} = \frac{1}{2}(\mu([k]) + \nu([k])) - \min(\mu, \nu)([k]), \quad (1)$$

where we use the notation $[k] = \{1, 2, \dots, k\}$. We also use the following partial order on measures on the set S : $\mu \preceq \nu$, if and only if for all $T \subseteq S$, $\mu(T) \leq \nu(T)$. The following observation is immediate from (1).

Observation 7. Fix $k \in \mathbb{N}$, $\varepsilon > 0$, and measures $\mu, \nu, \mu', \nu' \in \mathcal{M}_k$, such that $\mu' \preceq \mu$ and $\nu' \preceq \nu$. If

$$\|\mu - \nu\|_{TV} \geq \frac{1}{2}(\mu([k]) + \nu([k])) - \varepsilon,$$

then

$$\|\mu' - \nu'\|_{TV} \geq \frac{1}{2}(\mu'([k]) + \nu'([k])) - \varepsilon.$$

We will require the following fact in the proof of Lemma 5.

Lemma 8. Consider $\delta \in (0, 1)$ and a finite subset $S \subseteq [0, \infty)$ such that

$$\delta \cdot (|S| - 1) \cdot \sum_{x \in S} x \geq \sum_{x, y \in S, x \neq y} \min(x, y). \quad (2)$$

Then there exists a set $T \subseteq S$, such that $\sum_{x \in T} x \geq \frac{1}{2} \sum_{x \in S} x$ and $|T| \leq \lceil \delta(|S| - 1) \rceil$.

Proof. Let $n = |S|$, and let $a_1 \geq \dots \geq a_n \geq 0$ be the elements of S in decreasing order. Then,

$$\sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n \min(a_i, a_j) = \sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n a_{\max(i,j)} = \sum_{i=1}^n 2(i-1)a_i.$$

Letting $k = \lceil \delta(|S| - 1) \rceil$, we have

$$\sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n \min(a_i, a_j) \geq \sum_{i=k+1}^n 2(i-1)a_i \geq 2k \sum_{i=k+1}^n a_i \geq 2\delta(|S| - 1) \sum_{i=k+1}^n a_i.$$

Combining this inequality and (2) implies that $\sum_{i=k+1}^n a_i \leq \frac{1}{2} \sum_{x \in S} x$, therefore $\sum_{i=1}^k a_i \geq \frac{1}{2} \sum_{x \in S} x$. Hence the set $T = \{a_1, \dots, a_k\}$ satisfies both conditions of the lemma. \square

Proof of Lemma 5. We will show that each of the following statements implies the next one.

- I) There exists an ε -unrelated set $\mathcal{S} \subseteq \mathcal{P}_k$ of size n .
- II) There exists an ε -unrelated set $\mathcal{S} \subseteq \mathcal{M}_k$ of size n such that
 - (a) for all $\mu \in \mathcal{S}$, $\mu([k]) \leq 1$;
 - (b) $\sum_{\mu \in \mathcal{S}} \mu([k]) \geq n/4$;
 - (c) $\sum_{\mu \in \mathcal{S}} |\text{supp}(\mu)| < (2\varepsilon n + 1)k$;
- III) There exists an ε -unrelated set $\mathcal{S} \subseteq \mathcal{M}_k$ of size at least $n/14$ such that
 - (a) for all $\mu \in \mathcal{S}$, $|\text{supp}(\mu)| < 14(2\varepsilon + \frac{1}{n})k$;
 - (b) for all $\mu \in \mathcal{S}$, we have $\mu([k]) \geq 1/8$;
- IV) There exists a set satisfying all the conditions of the lemma.

For ease of notation, given a subset $\mathcal{S} \subseteq \mathcal{M}_k$, we define,

$$\Delta_{\mathcal{S}} = \sum_{\mu, \nu \in \mathcal{S}, \mu \neq \nu} \min(\mu, \nu).$$

Note that, if for some $\varepsilon \in [0, 1]$, $\mathcal{S} \subseteq \mathcal{P}_k$ is ε -unrelated, then (1) implies that

$$\begin{aligned} \Delta_{\mathcal{S}}([k]) &\leq \sum_{\mu, \nu \in \mathcal{S}, \mu \neq \nu} \frac{1}{2}(\mu([k]) + \nu([k])) - \|\mu - \nu\|_{TV} \\ &\leq \sum_{\mu, \nu \in \mathcal{S}, \mu \neq \nu} (1 - (1 - \varepsilon)) \\ &= \varepsilon |\mathcal{S}| \cdot (|\mathcal{S}| - 1). \end{aligned} \tag{3}$$

I \Rightarrow II: Suppose that $\mathcal{S}_I \subseteq \mathcal{P}_k$ is ε -unrelated, and let X be a random variable with state space $\{1, \dots, k\}$ such that

$$\mathbb{P}(X = i) = \frac{\sum_{\mu \in \mathcal{S}_I} \mu(\{i\})}{|\mathcal{S}_I|}.$$

We have

$$\mathbb{E} \left[\frac{\Delta_{\mathcal{S}_I}(X)}{\sum_{\mu \in \mathcal{S}_I} \mu(X)} \right] = \frac{1}{|\mathcal{S}_I|} \sum_{i=1}^k \Delta_{\mathcal{S}_I}(\{i\}) = \frac{1}{|\mathcal{S}_I|} \Delta_{\mathcal{S}_I}([k]) \stackrel{(3)}{\leq} \varepsilon(|\mathcal{S}_I| - 1),$$

Markov's inequality implies that

$$\mathbb{P} \left(\frac{\Delta_{\mathcal{S}_I}(X)}{\sum_{\mu \in \mathcal{S}_I} \mu(X)} \leq 2\varepsilon(|\mathcal{S}_I| - 1) \right) \geq \frac{1}{2}.$$

So if we let

$$A = \left\{ i : \frac{\Delta_{\mathcal{S}_I}(\{i\})}{\sum_{\mu \in \mathcal{S}_I} \mu(\{i\})} \leq 2\varepsilon(|\mathcal{S}_I| - 1) \right\},$$

then we have

$$\frac{1}{|\mathcal{S}_I|} \sum_{\mu \in \mathcal{S}_I} \mu(A) = \sum_{\mu \in \mathcal{S}_I} \sum_{i \in A} \mathbb{P}(X = i) \geq \frac{1}{2}. \quad (4)$$

By Lemma 8, for all $i \in A$ there exists a set $W_i \subseteq \mathcal{S}_I$ such that $|W_i| \leq \lceil 2\varepsilon(|\mathcal{S}_I| - 1) \rceil$, and

$$\sum_{\mu \in W_i} \mu(\{i\}) \geq \frac{1}{2} \sum_{\mu \in \mathcal{S}_I} \mu(\{i\}). \quad (5)$$

For $\mu \in \mathcal{S}_I$, let $Y_\mu = \{i : \mu \in W_i\}$. For any $Y \subseteq S$, define $R_Y : 2^S \rightarrow 2^S$ by $R_Y(T) = T \cap Y$ on elements of Y and zero elsewhere.

Let $\mathcal{S}_{II} = \{\mu \circ R_{Y_\mu}\}_{\mu \in \mathcal{S}_I}$. Since $\mu \circ R_{Y_\mu} \preceq \mu$, \mathcal{S}_{II} satisfies II(a). Furthermore, Observation 7 implies that the collection $\{\mu \circ R_{Y_\mu}\}_{\mu \in \mathcal{S}_I}$ is ε -unrelated. Furthermore, \mathcal{S}_{II} satisfies II(b) because

$$\sum_{\mu \in \mathcal{S}_I} \mu(Y_\mu) \stackrel{(5)}{\geq} \frac{1}{2} \sum_{i \in A} \sum_{\mu \in \mathcal{S}_I} \mu(\{i\}) = \frac{1}{2} \sum_{\mu \in \mathcal{S}_I} \mu(A) \stackrel{(4)}{\geq} \frac{1}{4} |\mathcal{S}_I|.$$

Finally, condition II(c) holds because

$$\sum_{\mu \in \mathcal{S}_I} |\text{supp}(\mu \circ R_{Y_\mu})| \leq \sum_{i \in A} |W_i| < 2\varepsilon(|\mathcal{S}_I| - 1)|A| + |A| \leq (2\varepsilon|\mathcal{S}_I| + 1)k.$$

II \Rightarrow III: Suppose that $\mathcal{S}_{II} \subseteq \mathcal{M}_k$ is an ε -unrelated collection of cardinality n satisfying all the conditions of II. We have $\max\{\mu([k])\}_{\mu \in \mathcal{S}_{II}} \leq 1$ and $\sum_{\mu \in \mathcal{S}_{II}} \mu([k]) \geq |\mathcal{S}_{II}|/4$. Therefore, there exists a subcollection $\mathcal{S}' \subseteq \mathcal{S}_{II}$ such that for all $\mu \in \mathcal{S}'$, we have $\mu([k]) \geq 1/8$, and

$$|\mathcal{S}'| \geq \left(\frac{1/4 - 1/8}{1 - 1/8} \right) |\mathcal{S}_{II}| \geq \frac{n}{7}.$$

By Markov's inequality, there exists a collection of measures \mathcal{S}_{III} such that $|\mathcal{S}_{\text{III}}| \geq \frac{1}{2}|\mathcal{S}'| \geq \frac{1}{14}|\mathcal{S}_{\text{II}}|$, where for all $\mu \in \mathcal{S}_{\text{III}}$,

$$\begin{aligned} \text{supp}(\mu) \leq 2 \frac{\sum_{\mu \in \mathcal{S}'} |\text{supp}(\mu)|}{|\mathcal{S}'|} &\leq 2 \frac{\sum_{\mu \in \mathcal{S}_{\text{II}}} |\text{supp}(\mu)|}{|\mathcal{S}'|} \\ &\leq 2 \frac{\sum_{\mu \in \mathcal{S}_{\text{II}}} |\text{supp}(\mu)|}{|\mathcal{S}_{\text{II}}|/7} \stackrel{\text{II(c)}}{\leq} 14k \left(2\varepsilon + \frac{1}{n}\right). \end{aligned}$$

The set \mathcal{S}_{III} has size at least $\frac{n}{14}$ and by construction satisfies conditions (a) and (b) of III.

III \Rightarrow IV: Suppose $\mathcal{S}_{\text{III}} \subseteq \mathcal{M}_k$ is a an ε -unrelated collection of cardinality at least $n/14$. For each measure $\mu \in \mathcal{S}_{\text{III}}$, let $Z_\mu \subseteq \{1, \dots, k\}$ be the set of $\lceil 16 \cdot \varepsilon (14k(2\varepsilon + \frac{1}{n})) \rceil$ elements of $\{1, \dots, k\}$ that has the largest measures with respect to μ (breaking ties arbitrarily). Since $\varepsilon \leq \frac{1}{8}$, for all $\mu \in \mathcal{S}_{\text{III}}$ we have

$$\mu(Z_\mu) \geq \frac{1}{8} \left(\frac{16 \cdot \varepsilon (14k(2\varepsilon + \frac{1}{n}))}{14k(2\varepsilon + \frac{1}{n})} \right) = 2\varepsilon. \quad (6)$$

Let $\mathcal{S}_{\text{IV}} = \left\{ \frac{\mu \circ R_{Z_\mu}}{\mu(Z_\mu)} : \mu \in \mathcal{S}_{\text{III}} \right\}$. Clearly $\mathcal{S}_{\text{IV}} \subseteq \mathcal{P}_k$, and $|\mathcal{S}_{\text{IV}}| \geq \frac{n}{14}$. Moreover, by our construction for all $\bar{\mu} \in \mathcal{S}_{\text{IV}}$, $|\text{supp}(\bar{\mu})| \leq \lceil 224\varepsilon(2\varepsilon + \frac{1}{n})k \rceil$. To complete the proof we need to show \mathcal{S}_{IV} is $\frac{1}{2}$ -unrelated. Note that if $\mu, \nu \in \mathcal{S}_{\text{III}}$, then Observation 7 implies that

$$\min(\nu \circ R_{Z_\nu}, \mu \circ R_{Z_\mu})([k]) \stackrel{(1)}{=} \frac{\mu(Z_\mu) + \nu(Z_\nu)}{2} - \|\mu \circ R_{Z_\nu} - \mu \circ R_{Z_\mu}\|_{TV} \leq \varepsilon.$$

Therefore,

$$\begin{aligned} \left\| \frac{\mu \circ R_{Z_\mu}}{\mu(Z_\mu)} - \frac{\nu \circ R_{Z_\nu}}{\nu(Z_\nu)} \right\|_{TV} &\stackrel{(1)}{=} 1 - \min \left(\frac{\mu \circ R_{Z_\mu}}{\mu(Z_\mu)}, \frac{\nu \circ R_{Z_\nu}}{\nu(Z_\nu)} \right)([k]) \\ &\stackrel{(6)}{\geq} 1 - \min \left(\frac{\mu \circ R_{Z_\mu}}{2\varepsilon}, \frac{\nu \circ R_{Z_\nu}}{2\varepsilon} \right)([k]) \stackrel{(1)}{\geq} \frac{1}{2}, \end{aligned}$$

completing the proof. \square

3. NEARLY EQUILATERAL SETS IN ℓ_1^d

We will need the following result of Kahane [10].

Theorem 9. *For every $\varepsilon \in (0, 1)$, there exists a mapping $K_\varepsilon : \mathbb{R} \rightarrow \ell_2^d$ such that $d \leq O(1/\varepsilon)$ and the following holds: For every $x, y \in \mathbb{R}$,*

$$(1 - \varepsilon)\sqrt{|x - y|} \leq \|K_\varepsilon(x) - K_\varepsilon(y)\|_2 \leq \sqrt{|x - y|}.$$

Proof of Theorem 2. Suppose that $f : U_n \rightarrow \ell_1^d$ is a $(1 + \varepsilon)$ -embedding scaled so that f is 1-Lipschitz. Consider the mapping $g : U_n \rightarrow \ell_2^{O(d/\varepsilon)}$ given by

$$g(x) = \left(K_\varepsilon(f(x)_i), K_\varepsilon(f_2(x)_i), \dots, K_\varepsilon(f_d(x)_i) \right),$$

where $f(x)_i$ denotes the i th coordinate of $f(x)$. By Theorem 9, for any $x, y \in U_n$, we have $\|g(x) - g(y)\|_2^2 \leq \|f(x) - f(y)\|_1$. On the other hand,

$$\|g(x) - g(y)\|_2^2 \geq (1 - \varepsilon)^2 \|f(x) - f(y)\|_1 \geq \frac{(1 - \varepsilon)^2}{1 + \varepsilon},$$

implying that $\|g(x) - g(y)\|_2 \geq (1 - \varepsilon)/\sqrt{1 + \varepsilon} \geq 1 - 2\varepsilon$. Thus g is a $(1 + 2\varepsilon)$ -embedding of U_n into $\ell_2^{O(d/\varepsilon)}$. But now by [2], for n sufficiently large, we have

$$d \geq \frac{\Omega(\log n)}{\varepsilon \log \frac{1}{\varepsilon}}.$$

□

4. EXTENSION TO k -ARY TREES

We now prove Corollary 3. Combining the next lemma with Theorem 1 yields the desired result.

Lemma 10. *For $h, k \geq 2$, let $B_{h,k}$ be a complete k -ary tree of height h . If $B_{h,k}$ admits a $(1 + \varepsilon)$ -embedding into ℓ_1^d for some $0 \leq \varepsilon \leq \frac{1}{8}$, then the $(1 + k^{\lceil h/2 \rceil})$ -star admits a $(1 + 4\varepsilon)$ -embedding into ℓ_1^d .*

Proof. Suppose that $f : B_{h,k} \rightarrow \ell_1^d$ is a $(1 + \varepsilon)$ -embedding. We may assume, without loss, that f is 1-Lipschitz. Letting $n = (1 + k^{\lceil h/2 \rceil})$, we construct an embedding $g : V_n \rightarrow \ell_1^d$ of the n -star as follows.

Let $r \in V_n$ denote the vertex of degree $n - 1$. We put $g(r) = 0$. Let S be the set of vertices in $B_{h,k}$ at height $\lceil h/2 \rceil$ (we use the convention that root has height zero). For any vertex $v \in S$, pick an arbitrary leaf x_v in the subtree rooted at v . Associate to every vertex $w \in V_n \setminus \{r\}$ a distinct element $\tilde{w} \in S$ and put

$$g(w) = \frac{f(x_{\tilde{w}}) - f(\tilde{w})}{h - \lceil h/2 \rceil}.$$

Since f is 1-Lipschitz, the same holds for g . Moreover for any two distinct elements $u, v \in S$ we have

$$\begin{aligned} 2(h - \lceil h/2 \rceil) + d_{B_{h,k}}(u, v) &= d_{B_{h,k}}(u, v) + d_{B_{h,k}}(x_v, v) + d_{B_{h,k}}(x_u, u) \\ &= d_{B_{h,k}}(x_u, x_v) \\ &\leq (1 + \varepsilon) \|f(x_u) - f(x_v)\|_1 \\ &\leq (1 + \varepsilon) \|(f(x_u) - f(u)) - (f(x_v) - f(v))\|_1 \\ &\quad + (1 + \varepsilon) \|f(u) - f(v)\|_1 \\ &\leq (1 + \varepsilon) \|(f(x_u) - f(u)) - (f(x_v) - f(v))\|_1 \\ &\quad + (1 + \varepsilon) d_{B_{h,k}}(u, v). \end{aligned}$$

Therefore,

$$\begin{aligned}
(1 + \varepsilon)\|(f(x_u) - f(u)) - (f(x_v) - f(v))\|_1 &\geq 2(h - \lceil h/2 \rceil) - \varepsilon d_{B_{h,k}}(u, v) \\
&\geq 2(h - \lceil h/2 \rceil) - 2\varepsilon \lceil h/2 \rceil \\
&\geq 2(h - \lceil h/2 \rceil) - 4\varepsilon(h - \lceil h/2 \rceil) \\
&\geq (2 - 4\varepsilon)(h - \lceil h/2 \rceil).
\end{aligned}$$

Since $\varepsilon \leq 1/8$, the preceding inequality bounds the distortion of g by $\frac{1+\varepsilon}{1-2\varepsilon} \leq 1 + 4\varepsilon$, completing the proof. \square

Acknowledgments. We are grateful to Jiri Matoušek for suggesting the approach of Section 3. This research was partially supported by NSF grant CCF-1217256.

REFERENCES

- [1] N. Alon and P. Pudlák. Equilateral sets in l_p^n . *Geom. Funct. Anal.*, 13(3):467–482, 2003.
- [2] Noga Alon. Problems and results in extremal combinatorics. I. *Discrete Math.*, 273(1-3):31–53, 2003. EuroComb’01 (Barcelona).
- [3] A. Andoni, M. Charikar, O. Neiman, and H. L. Nguyen. Near linear lower bound for dimension reduction in L_1 . To appear, *Proceedings of the 52nd Annual IEEE Conference on Foundations of Computer Science*, 2011.
- [4] Joshua D. Batson, Daniel A. Spielman, and Nikhil Srivastava. Twice-ramanujan sparsifiers. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing*, pages 255–262, 2009.
- [5] J. Bourgain, J. Lindenstrauss, and V. Milman. Approximation of zonoids by zonotopes. *Acta Math.*, 162(1-2):73–141, 1989.
- [6] Bo Brinkman and Moses Charikar. On the impossibility of dimension reduction in ℓ_1 . *J. ACM*, 52(5):766–788, 2005.
- [7] Moses Charikar and Amit Sahai. Dimension reduction in the ℓ_1 norm. In *43rd Annual Symposium on Foundations of Computer Science*, 2002.
- [8] P. Delsarte. An algebraic approach to the association schemes of coding theory. *Philips Res. Rep. Suppl.*, (10):vi+97, 1973.
- [9] William B. Johnson and Joram Lindenstrauss. Extensions of Lipschitz mappings into a Hilbert space. In *Conference in modern analysis and probability (New Haven, Conn., 1982)*, volume 26 of *Contemp. Math.*, pages 189–206. Amer. Math. Soc., Providence, RI, 1984.
- [10] Jean-Pierre Kahane. Hélices et quasi-hélices. In *Mathematical analysis and applications, Part B*, volume 7 of *Adv. in Math. Suppl. Stud.*, pages 417–433. Academic Press, New York, 1981.
- [11] J. R. Lee, A. de Mesmay, and M. Moharrami. Dimension reduction for finite trees in ℓ_1 . Prelim. version in SODA 2012. Preprint: arXiv:1108.2290, 2011.
- [12] J. R. Lee and A. Naor. Embedding the diamond graph in L_p and dimension reduction in L_1 . *Geom. Funct. Anal.*, 14(4):745–747, 2004.
- [13] J. Matoušek. Open problems on low-distortion embeddings of finite metric spaces. Online: <http://kam.mff.cuni.cz/~matousek/metrop.ps>.
- [14] Robert J. McEliece, Eugene R. Rodemich, Howard Rumsey, Jr., and Lloyd R. Welch. New upper bounds on the rate of a code via the Delsarte-MacWilliams inequalities. *IEEE Trans. Information Theory*, IT-23(2):157–166, 1977.

- [15] Ilan Newman and Yuri Rabinovich. On cut dimension of ℓ_1 metrics and volumes, and related sparsification techniques. *CoRR*, abs/1002.3541, 2010. Prelim. version in SODA 2012.
- [16] Oded Regev. Entropy-based bounds on dimension reduction in L_1 . arXiv:1108.1283, 2011.
- [17] Gideon Schechtman. More on embedding subspaces of L_p in l_r^n . *Compositio Math.*, 61(2):159–169, 1987.
- [18] Michel Talagrand. Embedding subspaces of L_1 into l_1^N . *Proc. Amer. Math. Soc.*, 108(2):363–369, 1990.

COMPUTER SCIENCE & ENGINEERING, UNIVERSITY OF WASHINGTON
E-mail address: jrl@cs.washington.edu

COMPUTER SCIENCE & ENGINEERING, UNIVERSITY OF WASHINGTON
E-mail address: mohammad@cs.washington.edu